Little Mathematics Library

A. I. FETISOV

PROOF IN GEOMETRY

Mir Publishers Moscow

First published 1977

На английском языке

CONTENTS

Foreword 7

1. What Is Proof? 8

2. Why Is Proof a Necessity? 12

3. What Should Be Meant by a Proof? 19

4. What Propositions May Be Accepted Without Proof? 44

FOREWORD

One fine day at the very start of a school year I happened to overhear two young girls chatting. They exchanged views on lessons, teachers, girl-friends, made remarks about new subjects. The elder was very much puzzled by lessons in geometry.

"Funny," she said, "the teacher enters the classroom, draws two equal triangles on the blackboard and next wastes the whole lesson proving to us that they are equal. I've no idea what's that for." "And how are you going to answer the lesson?" asked the younger. "I'll learn from the textbook although it's going to be a hard task trying to remember where every letter goes."

The same evening I heard that girl diligently studying geometry sitting at the window: "To prove the point let's superpose triangle $A'B'C'$ on triangle ABC superpose triangle $A'B'C'$ on triangle ABC she repeated time and again. Unfortunately, I do not know how well the girl did in geometry, but I should think the subject was not an easy one for her.

Some days later another pupil, Tolya, came to visit me, and he, too, had misgivings about geometry. Their teacher explained the theorem to the effect that an exterior angle of a triangle is greater than any of the interior angles not adjacent to it and made them learn the theorem at home. Tolya showed me a drawing from a textbook (Fig. 1) and asked whether there was any sense in a lengthy and complicated proof when the drawing showed quite clearly that the exterior angle of the triangle was obtuse and the interior angles not adjacent to it were acute. "But an obtuse angle," insisted Tolya, "is always greater than any acute angle. This is clear without proof." And I had to explain to Tolya that the point was by no means self-evident, and that there was every reason to insist on it being proved.

Quite recently a schoolboy showed me his test paper the mark for which, as he would have it, had been unjustly discounted. The problem dealt with an isosceles trapezoid with bases of 9 and 25 cm and with a side of 17 cm, it being required to find the altitude. To solve the problem a circle had been inscribed in the trapezoid and it was said that on the basis of the theorem on a circumscribed quadrilaterals (the sums of the opposite sides of a circumscribed quadrilateral are equal) one can inscribe a circle in the trapezoid ($9 + 25 = 17 + 17$). Next the altitude was identified with the diameter of the circle inscribed in the isosceles trapezoid which is equal to the geometrical mean of its bases

(the pupils proved that point in one of the problems solved earlier).

The solution had the appearance of being very simple and convincing, the teacher, however, pointed out that the reference to the theorem on a circumscribed quadrilateral had been incorrect. The boy was puzzled. "Isn't it true that the sums of opposite sides of a circumscribed quadrangle are equal? The sum of the bases of our trapezoid is equal to the sum of its sides, so a circle may be inscribed in it. What's wrong with that?"

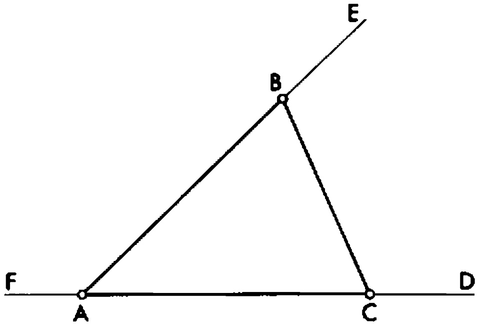

Fig.

One can cite many facts of the sort I have just been telling about. The pupils often fail to understand why truths should be proved that seem quite evident without proof, the proofs often appearing to be excessively complicated and cumbersome. It sometimes happens, too, that a seemingly clear and convincing proof turns out, upon closer scrutiny, to be incorrect.

This booklet was written with the aim of helping pupils clear up the following points:

1. What is proof?
2. What purpose does a proof serve?
3. What form should a proof take?
4. What may be accepted without proof in geometry?

§ 1. What Is Proof?

1. So let's ask ourselves: what is proof? Suppose you are trying to convince your opponent that the Earth has the shape of a sphere. You tell him about the horizon widening as the observer rises above the Earth's surface, about round-the-world trips, about

a disc-shaped shadow that falls from the Earth on the Moon in times of Lunar eclipses, etc.

Each of such statements designed to convince your opponent is termed an *argument* of the proof. What determines the strength or the convincibility of an argument? Let's discuss the last of the arguments cited above. We insist that the Earth must be round because its shadow is round. This statement is based on the fact that people know from experience that the shadow from all spherical bodies is round, and that, vice versa, a circular shadow is cast by spherical bodies irrespective of the position of a body. Thus, in this case, we first make use of the *facts* of our everyday experience concerning the properties of bodies belonging to the material world around us.

Next we draw a *conclusion* which in this case takes roughly the following form. "All the bodies that irrespective of their position cast a circular shadow are spherical." "At times of Lunar eclipses the Earth always casts a circular shadow on the Moon despite varying position it occupies relative to it." Hence, the conclusion: "The Earth is spherical."

Let's cite an example from physics.

The English physicist Maxwell in the sixties of the last century came to the conclusion that the velocity of propagation of electromagnetic oscillations through space is the same as that of light. This led him to the hypothesis that light, too, is a form of electromagnetic oscillations. To prove his hypothesis he should have made certain that the identity in properties of light and electromagnetic oscillations was not limited to the velocity of propagation, he should have provided the necessary arguments proving that the nature of both phenomena was the same. Such arguments were to come from the results of polarization experiments and several other facts which showed beyond doubt that the nature of optical and of electromagnetic oscillations was the same.

Let's cite, in addition, an arithmetical example. Let's take some odd numbers, square each of them and subtract unity from each of the squares thus obtained, e. g.:

$$7^2 - 1 = 48;\ 11^2 - 1 = 120;\ 5^2 - 1 = 24;$$
$$9^2 - 1 = 80;\ 15^2 - 1 = 224$$

etc. Looking at the numbers obtained in this way we note that they possess one common property, i. e. each of them can be divided by 8 without a remainder. After trying out several other odd numbers with identical results we should be prepared to state

the following hypothesis: "The square of every odd number minus unity is an integer multiple of 8."

Since we are now dealing with *any* odd number we should, in order to prove it, provide arguments which would do for *every* odd number. With this in mind, let's remember that every odd number is of the form $2n - 1$, where n is an arbitrary natural number. The square of an odd number minus unity may be written in the form $(2n - 1)^2 - 1$. Opening the brackets we obtain $(2n - 1)^2 - 1 = 4n^2 - 4n + 1 - 1 = 4n^2 - 4n = 4n(n - 1)$. The expression obtained is divisible by 8 for every natural n. Indeed, the multiplier 4 shows that the number $4n(n - 1)$ is divisible by 4. Moreover, $n - 1$ and n are two consecutive natural numbers, one of which is perforce even. Consequently, our expression must contain the multiplier 2 as well.

Hence, the number $4n(n - 1)$ is always an integer multiple of 8, and this is what we had to prove.

These examples will help us to understand the principal ways we take to gain knowledge about the world around us, its objects, its phenomena and the laws that govern them. The first way consists in carrying out numerous observations and experiments with objects and phenomena and in establishing on this basis the laws governing them. The examples cited above show that observations made it possible for people to establish the relationship between the shape of the body and its shadow; numerous experiments and observations confirmed the hypothesis about the electromagnetic nature of light; lastly, experiments which we carried out with the squares of odd numbers helped us to find out the property of such squares minus unity. This way – the establishment of general conclusions from observation of numerous specific cases – is termed *induction* (from the Latin word *inductio* – specific cases induce us to presume the existence of general relationships).

We take the alternative way when we are aware of some general laws and apply this knowledge to specific cases. This way is termed *deduction* (from the Latin word *deductio*). That was how in the last example we applied general rules of arithmetic to a specific problem, to the proof of the existence of some property common to all odd numbers.

This example shows that induction and deduction cannot be separated. The unity of induction and deduction is characteristic of scientific thinking.

It may easily be seen that in the process of any proof we make use of both ways. In search of arguments to prove some

proposition we turn to experience, to observations, to facts or to established propositions that have already been proven. On the basis of results thus obtained we draw a conclusion as to the validity, or falsity, of the proposition being proved.

2. Let's, however, return to geometry. Geometry studies spatial relationships of the material world. The term "spatial" is applied to such properties which determine the shape, the size and the relative position of objects. Evidently, the need of such knowledge springs from practical requirements of mankind: people have to measure lengths, areas and volumes to be able to design machines, to erect buildings, to build roads, canals, etc. Naturally, geometrical knowledge was initially obtained by way of induction from a very great number of observations and experiments. However, as geometrical facts accumulated, it became evident that many of them may be obtained from other facts by way of reasoning, e. by deduction, making special experiments unnecessary.

Thus, numerous observations and long experience convince us that "one and only one straight line passes through any two points" This fact enables us to state without any further experiment that "two different straight lines may not have more than one point in common" This new fact is obtained by very simple reasoning. Indeed, if we assume that two different straight lines have two common points we shall have to conclude that two different straight lines may pass through two points, and this contradicts the fact established earlier.

In the course of their practical activities men established a very great number of geometrical properties that reflect our knowledge of the spatial relationships of the material world. Careful studies of these properties showed that some of them may be obtained from the others as logical conclusions. This led to the idea of choosing from the whole lot of geometrical facts some of the most simple and general ones that could be accepted without proof and using them to deduce from them the rest of geometrical properties and relationships.

This idea appealed already to the geometers of ancient Greece, and they began to systematize geometrical facts known to them by deducing them from comparatively few fundamental propositions. Some 300 years B. C. Euclid of Alexandria made the most perfect outline of the geometry of his time. The outline included selective propositions which were accepted without proof, the so-called *axioms* (the Greek word αξιοζ means "worthy", "trustworthy"). Other propositions whose validity was tested by proof became

known as *theorems* (from the Greek word ϑεορεο – to think, to ponder).

The Euclidean geometry lived through many centuries, and even now the teaching of geometry at school in many aspects bears the marks of Euclid. Thus, in geometry we have comparatively few fundamental assumptions – axioms – obtained by means of induction and accepted without proof, the remaining geometrical facts being deduced from these by means of deductive reasoning. For this reason geometry is mainly a deductive science.

At present many geometers strive to reveal all the axioms necessary to build the geometrical system, keeping their number down to the minimum. This work has begun already in the last century and although much has already been accomplished it may not even now be regarded as complete.

In summing up this section we are now able to answer the question: what is proof in geometry? As we have seen, proof is a system of conclusions with the aid of which the validity of the proposition being proved is deduced from axioms and other propositions that have been proved before.

One question still remains: what is the guarantee of the truth of the propositions obtained by means of deductive reasoning?

The truth of a deduced conclusion stems from the fact that in it we apply some general laws to specific cases for it is absolutely obvious that something that is generally and always valid will remain valid in a specific case.

If, for instance, I say that the sum of the angles of every triangle is 180° and that ABC is a triangle there can be no doubt that $\angle A + \angle B + \angle C = 180^\circ$

If you study geometry carefully you will easily find out that that is exactly the way we reason in each case.

§ 2. Why Is Proof a Necessity?

1. Let's now try to answer the question: why is proof a necessity?

The need for proof follows from one of the fundamental laws of logic (logic is the science that deals with the laws of correct thinking) – *the law of sufficient reason*. This law includes the requirement that every statement made by us should be founded, i. e. that it should be accompanied by sufficiently strong arguments capable of upholding the truth of our statement, testifying to its compliance with the facts, with reality. Such arguments may consist

either in a reference to observation and experiment by means of which the statement could be verified, or in a correct reasoning made up of a system of judgements.

The argumentation of the latter type is most common in mathematics.

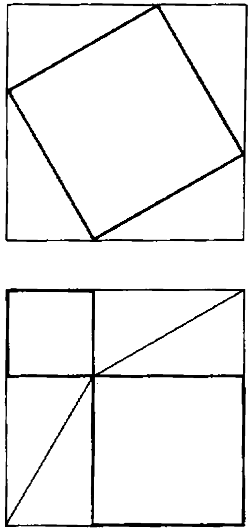

Fig.

Proof of a geometrical proposition aims at establishing its validity by means of logical deduction from facts known or proven before.

However, still the question springs up: should one bother about proof when the proposition to be proved is quite evident by itself?

This was the view taken by Indian mathematicians of the Middle Ages. They did not prove many geometrical propositions, but instead supplied them with expressive drawing with a single word "Look!" written above. Thus, for instance, the Pythagorean theorem appears in the book *Lilawaty* by the Indian mathematician Bhaskar Acharya in the following form (Fig. 2). The reader is expected to "see" from these two drawings that the sum of the areas of squares built on the legs of a right triangle is equal to the area of the square built on the hypotenuse.

Should we say that there is no proof in this case? Of course.

not. Should the reader just look at the drawing without pondering over it he could hardly be expected to arrive at any conclusion. The author actually presumes that the reader not only looks at the drawing, but thinks about it as well. The reader should understand that he has equal squares with equal areas before him. The first square is made up of four equal right-angled triangles and a square built on the hypotenuse and the second, of four identical right triangles and of two squares built on the legs. It remains to be realized that if we subtract equal quantities (the areas of four right-angled triangles) from equal quantities (the areas of two large equal squares) we shall obtain equal areas: the square built on the hypotenuse in the first instance and the two squares built on the legs in the second.

Still, however, aren't there theorems in geometry so obvious that no proof whatever is needed?

It is appropriate here to remark that an exact science cannot bear systematic recourse to the obvious for the concept of the obvious is very vague and unstable: what one person accepts as obvious, another may have very much in doubt. One should only recall the discrepancies in the testimonies of the eyewitnesses and the fact that it is sometimes very hard to arrive at the truth on the basis of such testimony.

An interesting geometrical example of a case when a seemingly obvious fact may be misleading may be cited. Here it is: I take a sheet of paper and draw on it a continuous closed line; next I take a pair of scissors and make a cut along this line. The question is: what will happen to the sheet of paper after the ends of the strip are stuck together? Presumably most of you will answer unhesitatingly: the sheet will be cut in two separate parts. This answer may, however, happen to be *wrong*. Let's make the following experiment: take a paper strip and paste its ends together to make a ring after giving it half a twist. We shall obtain the so-called Möbius strip (Fig. 3). (Möbius was a German mathematician who studied surfaces of that kind.) Should we now cut this strip along a closed line at approximately equal distances from both fringes the strip *would not be cut* in two separate parts – we should still have *one* strip. Facts of this sort make us think twice before relying on "obvious" considerations.

2. Let's discuss this point in more detail. Let's take for the first example the case of the schoolgirl mentioned above. The girl was puzzled when she saw the teacher draw two equal triangles and then heard her proving the seemingly obvious fact of their equality. Things actually took a quite different turn:

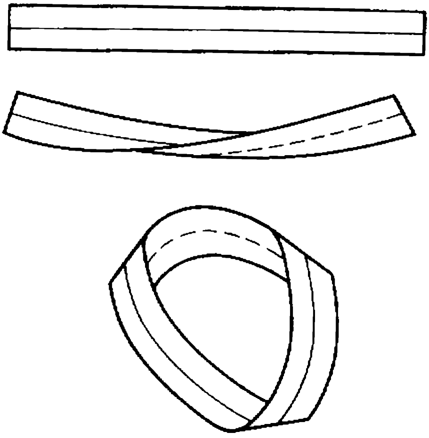

Fig. 3

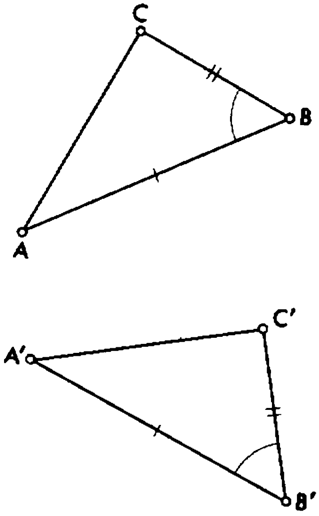

Fig. 4

the teacher did by no means draw two equal triangles, but, having drawn the triangle ABC (Fig. 4), said that the second triangle $A'B'C'$ was built so that $A'B' = AB$, $B'C' = BC$ and $\angle B = \angle B'$ and that we do not know whether $\angle A$ and $\angle A'$, $\angle C$ and $\angle C'$ and the sides $A'C'$ and AC are equal (for she did not build the angles $\angle A'$ and $\angle C'$ to be equal to the angles $\angle A$ and $\angle C$, respectively, and she did not make the side $A'C'$ equal to the side AC).

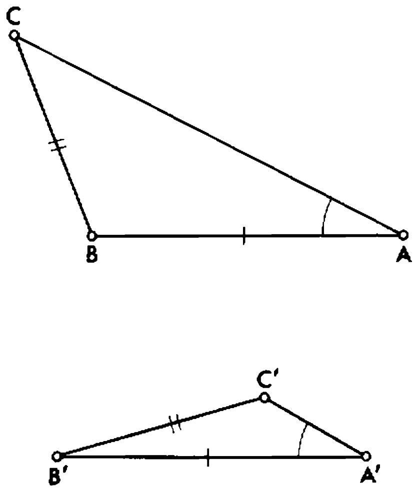

Fig. 5

Thus, in this case it is up to us to *deduce* the equality of the triangles, i. e. the equality of all their elements, from the conditions $A'B' = AB$; $B'C' = BC$ and $\angle B' = \angle B$, and this requires some consideration, i. e. requires proof.

It may easily be shown, too, that the equality of triangles based on the equality of three pairs of their respective elements is not at all so "obvious" as would appear at first glance. Let's modify the conditions of the first theorem: let two sides of one triangle be equal to two respective sides of another, let the angles be equal, too, though not the angles between these sides, but those lying opposite one of the equal sides, say, BC and $B'C'$ Let's write this condition for $\triangle ABC$ and $\triangle A'B'C'$: $A'B' = AB, B'C' = BC$ and $\angle A' = \angle A$. What is to be said about these triangles? By analogy with the first instance of equality of two triangles we could expect these triangles to be equal, too, but Fig. 5 convinces us that the triangles ABC and $A'B'C'$ drawn

in it are by no means equal although they satisfy the conditions $A'B' = AB$, $B'C' = BC$ and $\angle A' = \angle A$.

Examples of this sort tend to make us very careful in our deliberations and show with sufficient clarity that only a correct proof can guarantee the validity of the propositions being advanced.

3. Consider now a second theorem, the theorem on the exterior angle of a triangle which puzzled Tolya. Indeed, the drawing contained in the approved textbook shows a triangle whose exterior angle is obtuse and the interior angles not adjacent to it are acute, what can easily be judged without any measurement. But does it follow from this that the theorem requires no proof? Of

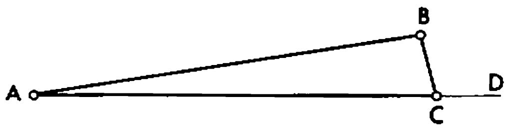

Fig. 6

course, not. For the theorem deals not only with the triangle drawn in the book, or, for that matter, on paper, on the blackboard, etc. but with any triangle whose shape may be quite unlike that shown in the textbook.

Let's imagine, for instance, that the point A moves away from the point C along a straight line. We shall then obtain the triangle ABC of the form shown in Fig. 6 where the angle at the point B will be obtuse, too. Should the point A move some ten metres away from the point C we would be unable to detect the difference between the interior and exterior angles with the aid of our school protractor. And should the point A move away from the point C by the distance, say, equal to that from the Earth to the Sun, one could say with absolute certainty that none of the existing instruments for measuring angles would be capable of detecting the difference between these angles. It follows that in the case of this theorem, too, we cannot say that it is "obvious" A rigorous proof of this theorem, however, *does not depend* on the specific shape of the triangle shown in the drawing and demonstrates that the theorem about the exterior angle of a triangle is valid for *all* triangles without exception, this not being dependent on the relative length of its sides. Therefore, even in cases when the difference between the interior and exterior angles so small as to defy detection with the aid of our instruments still are sure that it *exists*. This is because we have proved

that always, in all cases, the exterior angle of a triangle is greater than any interior angle not adjacent to it.

In this connection it is appropriate to look at the part played by the drawing in the proof of a geometrical theorem. One should keep in mind that the drawing is but an *auxiliary device* in the proof of the theorem, that it is only an *example*, only a *specific case* from a whole class of geometrical figures for which the theorem is being proved. For this reason it is very important to be able to distinguish the general and stable properties of the figure shown in the drawing from the specific and casual ones. For instance, the fact that the drawing in the approved textbook accompanying the theorem about the exterior angle of a triangle shows an obtuse exterior angle and acute interior angles is a mere coincidence. Obviously, it is not permissible to base the proof of a property common to all triangles on such casual facts.

An essential characteristic of a geometrical proof that in a great degree determines its necessity is the one that enables it to be used to establish *general* properties of spatial figures. If the inference was correct and was based on correct initial propositions, we may rest assured that the proposition we have proven is valid. Just because of this we are confident that every geometrical theorem, for instance, the Pythagorean theorem, is valid for a triangle of arbitrary size with the length of its sides varying from several millimetres to millions of kilometres.

4. There is, however, one more extremely important reason for the necessity of proof. It boils down to the fact that geometry is not a casual agglomeration of facts describing the spatial properties of bodies, but a *scientific system* built in accordance with rigorous laws. Within this system every theorem is structurally related to the totality of propositions established previously, and this relationship is brought to the surface by means of proof. For example, the proof of the well-known theorem on the sum of the interior angles of a triangle being equal to 180° is based on the properties of parallel lines and this points to a relationship existing between the theory of parallel lines and the properties of the sums of interior angles of polygons. In the same way the theory of similarity of figures as a whole is based upon the properties of parallel lines.

Thus, every geometrical theorem is connected with theorems proven before by a veritable system of reasonings, the same being true of the connections existing between the latter and the theorems proven still earlier and so on, the network of such

reasonings continuing down to the fundamental definitions and axioms that make up the corner-stones of the whole geometrical structure. This system of connections may be easily followed if one takes any geometrical theorem and considers all the propositions it is based on.

Summing up we may state the case for the necessity of proof as follows:

(a) In geometry only a few basic propositions – axioms – are accepted without proof. Other propositions – theorems – are subject to proof on the basis of these axioms with the aid of a set of judgements. The validity of the axioms themselves is guaranteed by the fact that they, as well as theorems based on them, have been verified by repeated observation and long-standing experience.

(b) The procedure of proof satisfies the requirement of one of the fundamental laws of human thinking – the law of sufficient reason that points to the necessity of rigorous argumentation to confirm the truth of our statements.

(c) A proof correctly constructed can be based only on propositions previously proved, no references to obvious fact being permitted.*

(d) Proof is also necessary to establish the general character of the proposition being proved, i. e. its applicability to all specific cases.

(e) Lastly, proofs help to line up geometrical facts into an elegant *system* of scientific knowledge, in which all interrelations between various properties of spatial forms are made tangible.

§ 3. What Should Be Meant by a Proof?

1. Let's turn now to the following question: what conditions should a proof satisfy for us to call it a *correct* one, i. e. one able to guarantee true conclusions from true assumptions? First of all note that every proof is made up of a series of judgements, therefore the validity, or falsity, of a proof depends on whether the corresponding judgements are correct, or erroneous.

As we have seen, deductive reasoning consists in the application of some general law to a specific case. To avoid an error in

---

* Many propositions of science previously considered unassailable because of their obvious character in due time turned out to be false. Every proposition of each science should be the object of rigorous proof.

the inference one should be aware of certain patterns with the aid of which the relations between all sorts of concepts, including those of geometry, are expressed. Let's show this with the aid of an example. Suppose we obtain the following inference: (1) The diagonals of all rectangles are equal. (2) All squares are rectangles. (3) Conclusion: the diagonals of all squares are equal.

What do we have in this case? The first proposition establishes some general law stating that all rectangles, i. e. a *whole class* of geometrical figures termed rectangles, belong to a class of quadrilaterals the diagonals of which are equal. The second

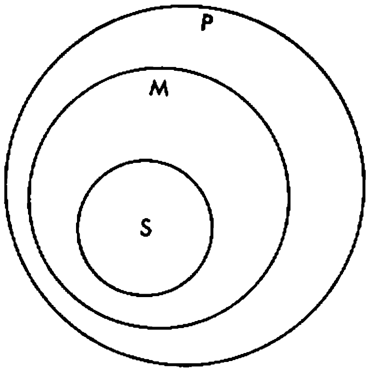

Fig. 7

proposition states that the entire class of squares is a part of the class of rectangles. Hence, we have every right to conclude that the entire class of squares is a part of the class of quadrilaterals having equal diagonals. Let's express this conclusion in a generalized form. Let's denote the widest class (quadrilaterals with equal diagonals) by the letter P, the intermediate class (rectangles) by the letter M, the smallest class (squares) by the letter S.

Then schematically our inference will take the following form:

(1) All M are P.

(2) All S are M.

(3) Conclusion: all S are P.

This relationship may easily be depicted graphically. Let's depict the largest class P by a large circle (Fig. 7). The class M will be depicted by a smaller circle lying entirely inside the first. Lastly, we shall depict the class S by the smallest circle placed inside the second circle. Obviously, with the circles placed as shown, the circle S will lie entirely inside the circle P.

This method of depicting the relationships between concepts, by the way, was proposed by the great mathematician Leonard Euler, Member of the St. Petersburg Academy of Sciences (1707-1783).

Such a pattern may be used to express other forms of judgement, as well. Consider now another inference that leads to a negative conclusion:

(1) All quadrilaterals whose sum of opposite angles is not equal to 180° cannot be inscribed in a circle.

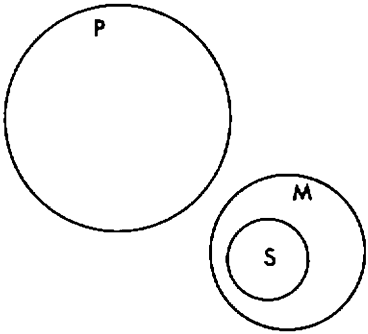

Fig. 8

(2) The sum of opposite angles of an oblique parallelogram is not equal to 180°

(3) Conclusion: an oblique parallelogram cannot be inscribed in a circle. Let's denote the class of quadrilaterals which cannot be inscribed in a circle by the letter P, the class of quadrilaterals whose sum of opposite angles is not equal to 180° by the letter M, the class of oblique parallelograms by the letter S. Then we shall find that our inference follows this pattern:

(1) None of the M's is P

(2) All S are M.

(3) Conclusion: none of the S's is P

This relationship, too, may be made quite visible with the aid of the Euler circles (Fig. 8).

The overwhelming majority of deductive inferences in geometry follows one or the other pattern.

2. Such depiction of relationships between geometrical concepts facilitates understanding of the structure of every judgement and the detection of an error in incorrect judgements.

By way of an example let's consider the reasoning of the pupil mentioned above which the teacher branded as erroneous. He obtained his inference in the following way:

(1) The sums of opposite sides of all circumscribed quadrilaterals are equal.

(2) The sums of opposite sides of the trapezoid under consideration are equal.

(3) Conclusion: the said trapezoid can be circumscribed about a circle.

Denoting the class of circumscribed quadrilaterals by *P*, the class of quadrilaterals having equal sums of opposite sides by *M*

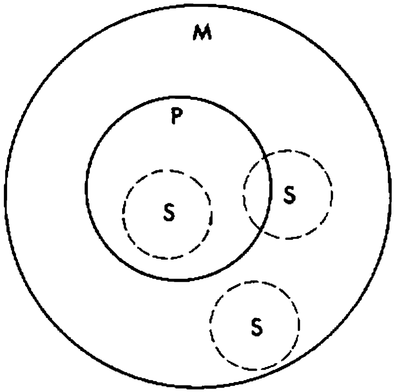

Fig. 9

and the class of trapezoids having the sum of bases equal to that of the sides by *S* we shall bring our inference in line with the following pattern:

(1) All *P* are *M*.

(2) All *S* are *M*.

(3) The conclusion that all *S* are *P* is *wrong* for using the Euler circles to depict the relationships between the classes (Fig. 9) we see that *P* and *S* lie inside *M*, but we are unable to draw any conclusion about the relationship between *S* and *P*

To make the error in the conclusion obtained above still more apparent let's cite as an example a quite similar inference:

(1) The sum of all adjacent angles is 180°

(2) The sum of two given angles is 180°

(3) Conclusion: therefore the given angles are adjacent. This of course, an erroneous conclusion for the sum of the

given angles may be 180° but they need not be adjacent (for instance, the opposite angles of an inscribed quadrilateral). How do such errors come about? The clue is that people making use of such reasoning refer to the *direct* theorem instead of to the *converse* of it. In the example with the circumscribed quadrilateral use has been made of the theorem stating that the sums of opposite sides of a circumscribed quadrilateral are aqual. However, the approved textbook does not contain the proof of the converse of that theorem to the effect that a circle can be inscribed in any quadrilateral with equal sums of opposite sides although such a proof is possible and will be presented below.

Should the theorem have been proved the correct judgement should follow the pattern:

(1) A circle can be inscribed in every quadrilateral with equal sums of opposite sides.

(2) The sum of the bases of the given trapezoid is equal to that of the sides.

(3) Conclusion: therefore, a circle can be inscribed in the given trapezoid. Naturally, this conclusion is quite correct for it has been constructed along the pattern shown in Fig. 6.

(1) All M are P

(2) All S are M.

(3) Conclusion: all S are P.

Thus, the mistake of the pupil was that he relied on the direct theorem instead of relying on the converse of it.

3. Let's prove this important converse theorem.

Theorem. *A circle can be inscribed in every quadrilateral with equal sums of the opposite sides.*

Note, to begin with, that if a circle can be inscribed in a quadrilateral, its centre will be equidistant from all its sides. Since the bisector is the locus of points that are equidistant from the sides of a quadrilateral the centre of the inscribed circle will lie on the bisector of each interior angle. Hence the centre of the inscribed circle is the point of intersection of the four bisectors of the interior angles of the quadrilateral.

Next, if at least three bisectors intersect at the same point, the fourth bisector will pass through that point, as well, and the said point is equidistant from all the four sides and is the centre of the inscribed circle. This can be proved by means of the same considerations that were used to prove the theorem on the existence of a circle inscribed in a triangle and we therefore leave it to the reader to prove it himself.

Now we shall turn to the main part of the proof. Suppose

we have a quadrilateral $ABCD$ (Fig. 10) for which the following relation holds:

$$AB + CD = BC + AD \qquad (1)$$

First of all we exclude the case when the given quadrangle turns out to be a rhombus, for rhombus's diagonals are the bisectors of its interior angles and because of this the point of their intersection is the centre of the inscribed circle, i. e. it is always possible to inscribe a circle in a rhombus. Therefore, let's suppose that two adjacent sides of our quadrangle are

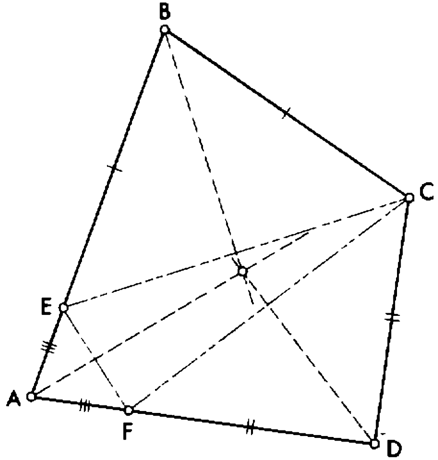

Fig. 10

unequal. Let, for instance, $AB > BC$. Then, as the result of equation (1) we shall have: $CD < AD$. Marking off the segment $BE = BC$ on AB we obtain an isosceles triangle BCE. Marking off the segment $DF = CD$ on AD we obtain an isosceles triangle CDF Let's prove that $\triangle AEF$ is isosceles, too. Indeed, let's transfer BC in equation (1) to the left and CD to the right and obtain: $AB - BC = AD - CD$. But $AB - BC = AE$, $AD - CD = AF$ Hence, $AE = AF$ and $\triangle AEF$ is an isosceles triangle. Now let's draw bisectors in three isosceles triangles thus obtained, i. e. the bisectors of $\angle B$, $\angle D$ and $\angle A$. These three bisectors are perpendicular to the bases CE, CF and EF and divide them in two. Hence they are perpendiculars erected from the mid-points of the sides of triangle CEF and must therefore intersect at one point. It follows that three bisectors of our quadrangle intersect at one point which,

as has been demonstrated above, is the centre of the inscribed circle.

4. Quite frequently one comes up against the following error in proof: instead of referring to the converse of a theorem people refer to the direct theorem. One must be very careful to avoid this error. For instance, when pupils are required to determine the type of the triangle with the sides of 3, 4 and 5 units of length one often hears that the triangle is right-angled because the sum of the squares of two of its sides, $3^2 + 4^2$, is equal to the square of the third, 5^2, reference being made to the Pythagorean theorem instead of to the converse of it. This converse theorem states that if the sum of the squares of two sides of a triangle is equal to the square of the third side, the triangle is right-angled.

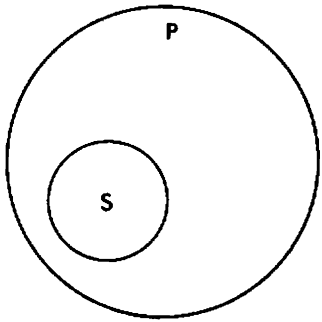

Fig. 11

Although the approved textbook does contain the proof of this theorem little attention is usually paid to it and this is the cause of the errors mentioned above.

In this connection it would be useful to determine the conditions under which both the direct and the converse theorems are true. We are already acquainted with examples when both a theorem and the converse of it hold, but one can cite as many examples when the theorem holds and the converse of it does not. For instance, a theorem states correctly that vertical angles are equal while the converse of it would have to contend that all equal angles are vertical ones, which is, of course, untrue.

To visualize the relationship between a theorem and the converse of it we shall again resort to a schematic representation of this relationship. If the theorem states: "All *S* are *P*" ("All pairs of angles vertical in respect to each other are pairs of equal angles"), the converse of it must contain the statement: "All *P* are *S*"

("All pairs of equal angles are pairs of angles vertical in respect to each other"). Representing the relationship between the concepts in the direct theorem with the aid of the Euler circles (Fig. 11) we shall see that the fact that the class *S* is a part of the class *P* generally enables us to contend only that "*Some P* are *S*" ("Some pairs of equal angles are pairs of angles vertical in respect of each other").

What are then the conditions for simultaneous validity of the proposition "All *S* are *P*" and the proposition "All *P* are *S*"?

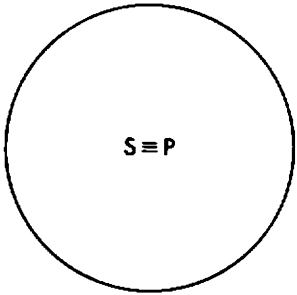

Fig. 12

It is quite obvious that this may happen if and only if the classes *S* and *P* are *identical* ($S \equiv P$). In this case the circle denoting *S* will coincide with the circle denoting *P* (Fig. 12). For instance, for the theorem "All isosceles triangles have equal angles adjacent to the base" the converse: "All triangles with equal angles at the base are isosceles triangles" holds as well. This is because the class of isosceles triangles and the class of triangles with equal angles at the base is one and the same class. In the same way the class of right triangles and that of the triangles whose square of one side is equal to the sum of squares of two other sides coincide. Our pupil was "lucky" to solve his problem despite the fact that he relied on the direct theorem instead of on the converse of it.

But this proved possible only because the class of quadrilaterals in which a circle can be inscribed coincides with the class of quadrilaterals whose sums of opposite sides are equal. (In this case both contentions "all *P* are *M*" and "all *M* are *P*" proved to be true – see p. 22.)

This investigation demonstrates at the same time that the converse of a theorem, should it prove true, is by no means an

obvious corollary of the direct theorem and *should always be the object of a special proof.*

5. It may sometimes appear that the direct theorem and the converse of it do not comply to the pattern "All S are P" and "All P are S" This happens when these theorems are expressed in the form of the so-called "conditional reasoning" which may be schematically written in the form: "If A is B, C is D." For example: "If a quadrilateral is circumscribed about a circle, the sums of its opposite sides will be equal." The first part of the sentence, "If A is B", is termed the *condition* of the theorem, and the second, "C is D", is termed its *conclusion*. When the converse theorem is derived from the direct one the conclusion and the condition change places. In many cases the conditional form of a theorem is more customary than the form "All S are P" which is termed the "categoric" form. However, it may easily be seen that the difference is inessential and that every conditional reasoning may easily be transformed into the categoric one, and vice versa. For example, the theorem expressed in the conditional form "If two parallel lines are intersected by a third line, the alternate interior angles will be equal" may be expressed in the categoric form: "Parallel lines intersected by a third line form equal alternate interior angles." Hence, our reasoning remains true of the theorems expressed in the conditional form, as well. Here, too, the simultaneous validity of the direct and the converse theorem is due to the fact that the classes of the respective concepts coincide. Thus, in the example considered above both the direct and the converse theorem hold, since the class of "parallel lines" is identical to the class of "the lines which, being intersected by a third line, form equal alternate interior angles"

6. Let's now turn to other defects of proof. Quite often the source of the error in proof is that specific cases are made the basis of the proof while other properties of the figure under consideration are overlooked. That was the mistake Tolya made in trying to prove the general theorem about the exterior angle of every triangle while limiting his discussion to the case of the acute triangle all the exterior angles of which are, indeed, obtuse while all the interior ones are acute.

Let's cite another example of a similar error in proof which this time is much less apparent. We have presented above the example of two *unequal* triangles (Fig. 4) whose two respective sides and an angle opposite one of the sides were, nevertheless, equal. Let's now present "proofs" that despite established facts the triangles satisfying the above conditions will necessarily be equal.

Another interesting point in connection with this proof is that it is very much like the proof of the third criterion of the equality of triangles in the approved textbook.

So let it be given that in $\triangle ABC$ and $\triangle A'B'C'$ (Fig. 13) $AB = A'B'$, $AC = A'C'$ and $\angle C = \angle C'$ To prove our point let's place $\triangle A'B'C'$ on $\triangle ABC$ so that side AB coincides with $A'B'$ and the point C' occupies the position C'' Let's connect the points C and C'' presuming that the segment CC'' will intersect the side AB between the points A and B (Fig. 13*a*). The condition

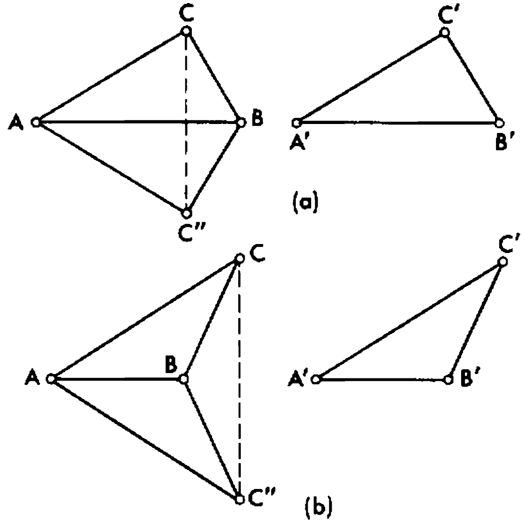

Fig. 13

stipulates that $\triangle ACC''$ is an isosceles triangle ($AC = AC''$) and consequently $\angle ACC'' = \angle AC''C$. Since $\angle C = \angle C''$ we find, after subtracting equal angles from equal angles, that $\angle BCC'' = \angle BC''C$, as well, and hence $\triangle CBC''$ will, too, be isosceles. Therefore, $BC = BC''$ and consequently $\triangle ABC = \triangle ABC''$ because all three of their sides are equal. Thus, $\triangle ABC = A'B'C'$

Should the segment CC'' intersect the line AB outside the segment AB, the theorem will still be valid (Fig. 13*b*). Indeed, $\triangle ACC''$ is in this case an isosceles one, as well, and $\angle ACC'' = \angle AC''C$. But since $\angle C = \angle C''$, after subtracting these angles from the angles of the previous equation we shall again find that $\angle BCC'' = \angle BC''C$ and that $\triangle BCC''$ is an isosceles one

with $BC = BC''$, and thus we again arrive at the third criterion of the equality of triangles, i. e. again $\triangle ABC = \triangle A'B'C'$

It seems that we have presented a sufficiently complete proof and exhausted all the possibilities. However, one more possibility was overlooked, i. e. that when the segment CC'' passes through the end of the segment AB. The segment CC'' in Fig. 14 passes through the point B. It may easily be seen that in that case our reasoning fails and the triangles may prove to be quite different, as shown in Fig. 14.

There is another very instructive example of an error of this sort, namely, the theorems on the lateral area of an oblique prism

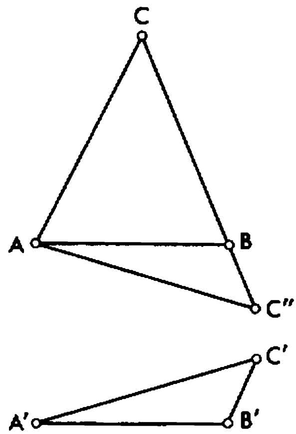

Fig. 14

and on the homogeneity of a right and an oblique prism. The first of these theorems states: "The lateral area of a prism is equal to the product of the perimeter of the normal section and the lateral edge." The second theorem states: "Every prism is homogeneous to the right prism whose base is the normal section of the oblique prism and whose altitude is its lateral edge." It is, however, easily seen that both theorems have in fact been proven only for a specific case, namely, that when the edges of the prism are long enough to enable a normal section to be drawn. At the same time there exists a whole class of prisms for which *it is impossible to draw a normal section which would intersect all the lateral edges.* These are extremely oblique prisms of very small altitude (Fig. 15). In such a prism a section perpendicular

to one of the lateral edges shall not intersect all the other edges, and the reasoning used to prove the stated propositions becomes inapplicable. In this case the source of the error lies in our habit of picturing the prism as a brick of sufficiently great height while at the same time short "table" prisms are practically never to be seen on a blackboard, in a notebook, or in a textbook. This example also shows how particular we must be with the drawing we use to illustrate the proof. Every time we make some construction we should ask ourselves: "Is this construction possible in every case?" Should this question have been asked when the above-mentioned proof of the propositions relating to an

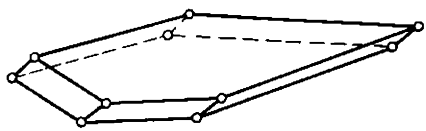

Fig. 15

oblique prism was conducted it would have been easy to find an example of a prism for which it is impossible to draw a normal section.

7. The essence of the error in the last two examples is that *the proof is effected not for the proposition to be proved, but for some specific case* relating to the peculiarities of the figure used in the process of proof. Another example of a similar error may be cited, this time, however, of a more subtle error and not so apparent.

The subject will be the proof of the existence of incommensurable segments that is usually presented in the school course of elementary geometry. Let's present a short reminder of the general course of reasoning leading to this proof. To begin with, a definition of a common unit of length of two segments is made wherein it is established that this unit of length is laid off a whole number of times along the sum and the difference of the given segments. Next, a method of finding the common unit of length is described of which already Euclid has been aware. The essence of that method is that the smaller segment is laid off on the greater one, next the first difference is laid off on the smaller segment, the second difference on the first difference, etc. The difference that when laid off on the preceding one leaves no difference is the greatest common unit of length of the two segments. A further definition states that the segments having a common unit of length

are termed *commensurable* and those that have no such unit of length are termed *incommensurable*. However, the very fact of the existence of incommensurable segments should have been proved by the discovery of at least one pair of such segments. The usual example cited is that of incommensurability of the diagonal and the side of a square. The proof is conducted on the basis of the Euclidean method of successive laying off first the side of the square on its diagonal, then the difference obtained on the side, etc. It turns out in the process that the difference between the diagonal and the side becomes the side of a new square that should be laid off on a new diagonal, etc., and that consequently such a process of successive laying off will never end and the greatest common unit of length of the diagonal and the side of a square cannot be found. Next the conclusion is drawn: therefore, a common unit of length of the diagonal of a square and its side cannot be found and the segments are incommensurable.

What is wrong with this conclusion? The error here lies in the fact that the *impossibility of finding a common unit of length by the use of the Euclidean method* in no way proves that such a common unit does not exist. For should we fail to find some object with the aid of a certain method it would not mean that it could not be found with other means. Under no circumstances would we be prepared to accept such a reasoning, for example: "Electrons are not visible in any microscope, therefore they do not exist." No doubt, it is easy to counter reasoning of this kind by such a remark: "There are other means and methods, besides the microscope, that we can use to detect the existence of the electrons."

To perfect the proof of the existence of incommensurable segments it is necessary to begin by proving the following proposition.

If the process of searching for the greatest common unit of length of two segments continues infinitely long such segments would be incommensurable.

Let's prove this important proposition.

Let $\bar{a}$ and $\bar{b}$ be the given segments (the lines above denote segments, the letters without the lines denote numbers) with $\bar{a} > \bar{b}$. Suppose we lay off successively $\bar{b}$ along $\bar{a}$, the first difference $\bar{r}_1$ along $\bar{b}$, etc. and obtain as a result an *unlimited* series of differences: $\bar{r}_1, \bar{r}_2, \bar{r}_3, \ldots$, every preceding segment being greater than the following. Thus, we shall have

$$\bar{a} > \bar{b} > \bar{r}_1 > \bar{r}_2 > \bar{r}_3 >$$

Suppose the segments $\bar{a}$ and $\bar{b}$ have a common unit of length p, and, by the property of a common unit of length, it is possible

to lay it off a whole number of times on $\bar{a}$, on $\bar{b}$ and on each of the differences $\bar{r}_1$, $\bar{r}_2$, $\bar{r}_3$, Suppose this unit of length can be laid off m times on $\bar{a}$, n times on $\bar{b}$, n_1 times on $\bar{r}_1$, n_2 times on $\bar{r}_2$, n_k times on $\bar{r}_k$, etc. The numbers m, n, n_1, n_2, n_3, are positive integers and because of the inequalities between the segments we shall have respective inequalities between these numbers:

$$m > n > n_1 > n_2 > n_3 >$$

Since we made the assumption that the series of segments continues indefinitely, the series m, n, n_1, n_2, n_3, must continue indefinitely, as well, but *this is impossible* because a series of continuously decreasing positive integers cannot be infinite. The resulting contradiction forces us to drop the assumption that there is a common unit of length of such segments and to admit that they are incommensurable. The example of the square demonstrates the existence of the segments for which the process of successive laying off will never end, therefore, the diagonal of the square is incommensurable with its side.

Without this additional proposition the proof of the incommensurability of the segments does not strike its point for it proves quite another proposition than the one we were required to prove.

8. Frequently another type of error springs up in the course of a proof. This occurs when reference is made to propositions that have not been proven before. It also happens, although not so often, that the person proving the theorem makes a reference just to the proposition he is trying to prove. For example, sometimes the following conversation between the teacher and the pupil may be heard: The teacher asks: "Why are these lines perpendicular?" The pupil answers: "Because the angle between them is right." "And why is it right?" "Because the lines are perpendicular."

Such an error is termed "a vicious circle in proof" and in so apparent a form is comparatively rare. More often one meets it in a subtle form. For example, the pupil was required to solve a problem: "Prove that if two bisectors of a triangle are equal it must be an isosceles triangle."

The proof was constructed as follows: "Let in $\triangle ABC$ the bisector AM be equal to bisector BN (Fig. 16). Consider $\triangle ABM$ and $\triangle ABN$ which are equal since $AM = BN$, AB is common to both and $\angle ABN = \angle BAM$, being the halves of equal angles at the bases. Hence, $\triangle ABM = \triangle ABN$ and therefore $AN = BM$. Next consider $\triangle ACM$ and $\triangle BCN$ which are equal since $AM = BN$ and since the respective angles adjoining these sides are equal, as well.

Therefore, $CN = CM$ and hence $AN + NC = BM + CM$, $AC = BC$, just what was to be proved."

The proof is erroneous because it contains the reference to the equality of angles at the base of the triangle, which is due to the fact that the triangle is isosceles, just the very proposition that had to be proved.

In some cases the proof is based on unproved propositions being regarded as obvious although the propositions are not part of axioms. Let's consider two examples. The problem of the mutual

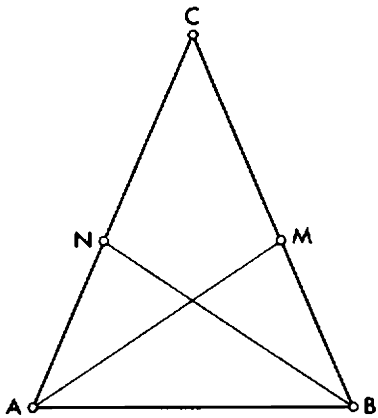

Fig. 16

position of a straight line and a circle is subdivided into three cases: (1) the distance between the line and the centre of the circle is greater than the radius – the line passes *outside* of the circle; (2) the distance from the line to the centre is equal to the radius – the line has one and only one common point with the circle (a tangent line); (3) the distance from the line to the centre is less than the radius – the line has two common points with the circle (a secant).

Note that the first two propositions are usually proved correctly while in the third case the agrument usually runs as follows: "The line passes through a point within the circle and in this case it *obviously* intersects the circle." It may easily be seen that the word "obviously" hides a very important geometrical proposition: "Every straight line that passes through an interior point of a circle intersects the circle." True, this proposition is rather obvious, but we discussed already how vague and indefinite the latter concept is.

Therefore, this proposition should either be proved on the basis of other propositions or made an axiom.

By way of the second example we shall cite the proof of the converse of the theorem on the circumscribed quadrangle which is contained in some courses of elementary geometry.

We are required to prove that *if the sums of opposite sides of a quadrangle are equal a circle can be inscribed in this quadrangle.*

We quote the proof: "Given $AB + CD = BC + AD$ (Fig. 17). Let's draw a circle which touches the sides AB, BC and CD. Let's prove that it will touch the side AD, too. Let's suppose the

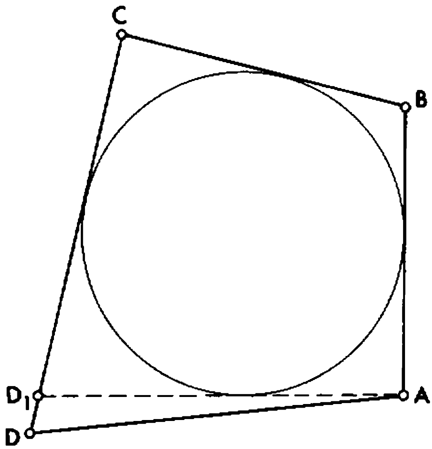

Fig. 17

circle failed to touch the side AD. Drawing the tangent AD_1 through the point A we shall obtain the circumscribed quadrangle $ABCD_1$ where, in compliance with the direct theorem, $AB + CD_1 = BC + AD_1$. Subtracting this equation term by term from the given equation we obtain $CD - CD_1 = AD_1 - AD$ or $DD_1 = AD_1 - AD$ which is impossible (the difference between two sides of $\triangle ADD_1$ cannot be equal to the third side). Therefore, the circle that is tangent to the sides AB, BC, and CD is tangent to the side AD, as well."

The error in this proof is that it is based on the knowledge of the position of the point A which has not, as yet, been gained in the proof: one should first prove that the *point of tangency of the circle lies between the points A and B*. Should the position of the points A and D be such as shown in Fig. 18

it would be impossible to apply the reasoning contained in this proof. It is quite possible to prove that the points of tangency must lie between A and B and between C and D, but this involves rather lengthy proceedings and because of that it is better to use the proof shown above (see p. 23).

Therefore, we should give the following answer to the question as to what form a proof should take to be correct, i. e. to be a guarantee of the truth of the proposition being proved:

(a) The proof should be based only on true propositions, i. e. on axioms and on the theorems that have been proved before.

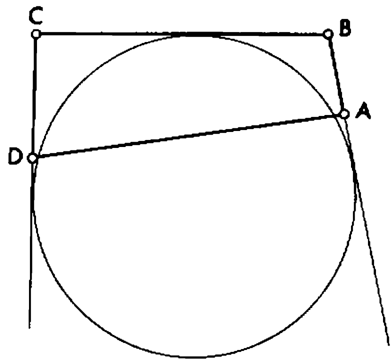

Fig. 18

(b) All inferences making up the proof should be correctly constructed.

(c) The aim of the proof, i. e. the establishment of the truth of the proposition being proved, should always be kept in mind and should not be substituted for by some other proposition.*

9. The need to fulfil these requirements poses a natural question: how can correct proofs be found?

Let's give some advice how to cope with this problem. When we are required to prove some geometrical proposition we should begin with clearly defining the *main idea* which should be the object of proof. Frequently this idea remains obscure. For example, we are offered to "prove that having connected successively the middle points of the sides of a quadrilateral we should obtain a parallelogram" Whereby are we to prove that we shall obtain

* As it was the case in example on p. 30.

a parallelogram? To answer the question we should recall the *definition* of the parallelogram as a quadrilateral the pairs of opposite sides of which are parallel. Hence, we must prove that the segments obtained shall be parallel.

After the proposition to be proved has been defined one should use the text of the theorem to define the conditions that are given there and that are needed for the proof. The example cited above states that we connect the *middle points* of the sides of a

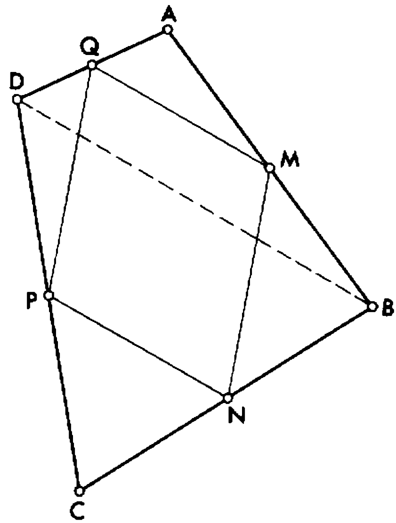

Fig. 19

quadrilateral – this means that a point is chosen on each side that divides it into two equal parts.

We write all this down in a symbolic form normally used in school practice and consisting of two sections headed "Given" and "To be proved" Thus, in our example if $ABCD$ is the quadrilateral (Fig. 19) and M, N, P, Q the middle points of its sides we can write our theorem in the form:

Given: in the quadrilateral $ABCD$ $MA = MB$, $NB = NC$, $PC = PD$, $QD = QA$.

To be proved: $MN \parallel PQ$, $MQ \parallel NP$

This notation is followed by the proof of the theorem. In the course of that proof we should make use of the axioms and theorems established before together (one should well remember it)

with such special relationships that may be stated in the conditions of the theorem.

10. But how should that set of judgements be found that would connect the proposition being proved with those established before and with the conditions of the theorem? How should a choice be made from among a multitude of sundry propositions of just those that would help us to prove our theorem?

The wisest way is to choose the proposition to be proved as the starting point of our quest and to formulate the problem as follows: the corollary of what proposition could it be? Should such a proposition be found and should it at the same time be the consequence of the conditions and theorems proved before our problem would be solved. Otherwise we are to formulate the same problem, but this time in respect of this new proposition, etc. Such a way of thinking in science is termed *analysis*.

In the example with the quadrilateral we are now considering we have to prove that some segments are parallel. At the same time we should remember that these segments connect the middle points of the quadrilateral's sides. Having established it we ask ourselves: is there a proposition among those proved before that deals with the parallelism of segments connecting the middle points of the sides of a polygon? One of such propositions is the theorem on the median of a triangle that says that the segment connecting the mid-points of two sides of the triangle is parallel to the third side and equal to half its length. But the figure under consideration has no such triangles. However, such a triangle can easily be constructed in it. Let's draw, for instance, the diagonal BD. This gives us at once two triangles ABD and BCD in which the segments MQ and NP act as medians. Hence, $MQ \parallel BD$ and $NP \parallel BD$ and, consequently, $NP \parallel MQ$. In the same way, after drawing the second diagonal we could prove that $MN \parallel PQ$. As it happens, such a construction is not necessary for from the first pair of triangles we have $MQ = {}^1/_2 BD$ and $NP = {}^1/_2 BD$ and hence $MQ = NP$, i. e. the opposite sides MQ and NP of the quadrilateral $MNPQ$ are not only parallel, but equal as well, and this gives us directly that the quadrilateral is a parallelogram.

For the second example let's take the well-known theorem on the sum of the interior angles of a triangle. In this case the text does not contain any special conditions and we therefore should write down only what is to be proved: in $\triangle ABC$ (Fig. 20) $\alpha + \beta + \gamma = 180°$

We see from the context of the proposition being proved that we shall have to add up the three interior angles of the triangle.

This addition is conveniently accomplished on the figure itself. Let's build the angle $\gamma' = \gamma$ at the vertex B of the angle β. Then the side BD of the angle γ' will be parallel to AC because the alternate angles made by the transversal BC are equal. Protracting the side AB beyond the point B we shall obtain $\angle CBE$ which we shall denote α', $\alpha' = \alpha$, both being the corresponding angles that the transversal AB makes with the two parallel lines. Thus we have $\alpha' + \beta + \gamma' = 180^\circ$ since together these angles make

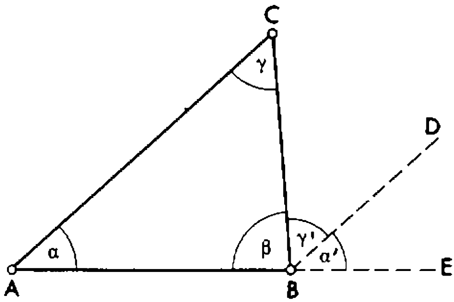

Fig. 20

up a straight angle. Hence, taking into account the equality of angles $\alpha' = \alpha$, $\gamma' = \gamma$ we obtain the required relationship

$$\alpha + \beta + \gamma = 180^\circ$$

In both examples cited above it did not take us long to find the necessary relationships. However, there are cases when such a relationship is established by means of a series of auxiliary propositions. In these cases the analysis becomes more lengthy and complicated.

11. Let's cite an example of a more complicated analysis. It is required to prove the following proposition: *If a circle is circumscribed about a triangle and perpendiculars to the sides of the triangle are dropped from an arbitrary point on it their bases will lie on a single line (the Simson line).*

Let's make the analysis. Let ABC be the given triangle (Fig. 21), M is the point on the circumscribed circle, N, P, Q are the respective projections of this point on the sides BC, CA, AB of the given triangle. It is required to prove that N, P and Q lie on the same line. We may write down the proposition to be proved noting that the conditions that the points N, P and Q are on the same

line is equivalent to the proposition that the angle NPQ is a straight one. Thus we have:

Given: $MN \perp BC$, $MP \perp CA$, $MQ \perp AB$. The point M is on the circle circumscribed about $\triangle ABC$.

To be proved: $\angle NPQ = 180°$ Considering the angle NPQ we see that it is made up of $\angle MPN = \delta$, $MPA = 90°$ and $\angle APQ = \alpha$. The proposition would have been proved had we been able to prove that $\angle NPQ = \delta + 90° + \alpha = 180°$ But to this end it is sufficient

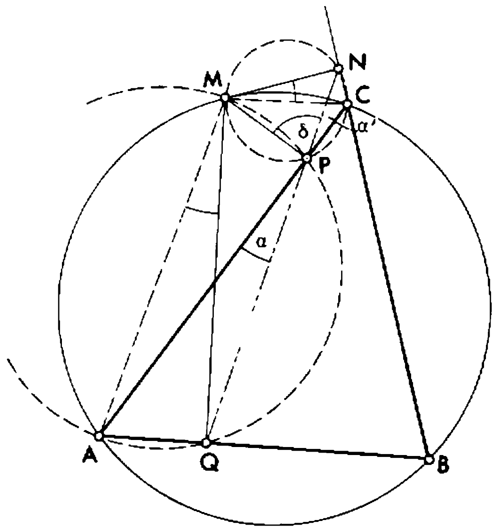

Fig. 21

to prove that $\alpha + \delta = 90°$ Consider $\angle CPN = \alpha'$. $\angle MPC = 90°$ and, by force of this, $\alpha' + \delta = 90°$ Thus, the theorem will have been proven if we were able to prove that $\alpha' = \alpha$. We shall try to establish the sought equation by considering new angles for which purpose we shall make use of the conditions of the theorem. The end points of the right angles APM and AQM rest upon the segment AM. Therefore a circle that would be built with AM as its diameter would pass through the points P and Q. Because of the properties of the inscribed angles $\angle AMQ = \angle APQ = \alpha$. In the same way, constructing a circle with the segment MC as its diameter we shall see that it will pass through P and N and in compliance with the properties of inscribed angles $\angle CMN = \angle CPN = \alpha'$. Let's try now to prove that $\angle AMQ = \angle CMN$. To this end note that the quadrilateral $ABCM$ is an inscribed one

and for this reason the sum of its opposite angles is 180°:

$$\left.\begin{aligned} \angle AMC + \angle B &= 180^\circ \\ \text{or} \qquad & \\ \angle AMQ + \angle QMC + \angle B &= 180^\circ \end{aligned}\right\} \qquad (1)$$

On the other hand, the angles at the points Q and N of the quadrilateral $BQMN$ are right angles, therefore the sum of its other two angles is 180°:

$$\left.\begin{aligned} \angle QMN + \angle B &= 180^\circ \\ \text{or} \qquad & \\ \angle QMC + \angle CMN + \angle B &= 180^\circ \end{aligned}\right\} \qquad (2)$$

Comparing equations (1) and (2) we obtain:

$$\angle QMC + \angle CMN + \angle B = \angle AMQ + \angle QMC + \angle B$$

whence

$$\angle CMN = \angle AMQ,$$

e. $\alpha' = \alpha$.

As we have already seen this results in $\alpha + \delta = 90^\circ$ and $\alpha + \delta + 90^\circ = 180^\circ$ and, lastly, $\angle NPQ = 180^\circ$

If we were to reconstruct the sequence of the proof we would have to move in the opposite direction: first we should have proved that $\angle AMQ = \angle CMN$; next we would have established the equalities

$$\angle AMQ = \angle MQN \text{ and } \angle CMN = \angle CPN$$

Lastly, making use of the fact that $\angle CPA = \angle CPN + \angle MPN + 90^\circ = 180^\circ$ we would have obtained that $\angle NPQ = \angle MPN + 90^\circ + \angle APQ = 180^\circ$, as well, i. e. that the points N, P and Q lie on the same line.

This method reciprocal to that of analysis is usually employed to prove theorems in textbooks and in class and is termed *synthesis*. It is easier and more natural to present the proof of a theorem by way of the synthetic method, but we should not forget that in *looking for* the proof we must perforce make use of analysis.

Thus, analysis and synthesis are two inseparable steps of the same process – that of constructing the proof of a given theorem. Analysis is the method of finding the proof, synthesis – that of presenting it.

It is, of course, not easy while looking for the proof of some proposition to establish the necessary sequence of judgements. It is not always possible to strike the right track at once, sometimes one has to drop the planned route and take another.

Here is an example. Suppose we have to prove the proposition: "If two medians of a triangle are equal the triangle is isosceles." Given: $\triangle ABC$ the medians AM and BN of which are equal. It may at first appear to be appropriate to consider the triangles ABM and ABN and to prove their equality. However, it is easily seen that we lack data for such a proof – we only know that $AM = BN$ and that the side AB is common to both triangles. We have neither the condition of equality of the angles, nor the condition of equality of the third sides. Therefore, we shall have to relinquish this way. Similarly, we shall come to the conclusion that it is senseless to consider the triangles ACM and BCN for we have not enough data to prove them equal either. Having dropped those two possibilities, let's look for some new ones. Let's denote the point of intersection of the two medians by P and consider the triangles ANP and BMP Since the medians are equal and since the point P lies on one-third of each median we shall find that $PN = PM$, $PA = PB$ and $\angle APN = \angle BPM$, both being vertical angles. Therefore, $\triangle ANP = \triangle BPM$ and consequently $AN = BM$. These segments being halves of the respective sides it follows that $AC = BC$, which is what was required to prove.

Analytical skill and the ability to find proofs independently are the result of repeated exercises and to this end one should systematically work on problems dealing with proofs.

12. To conclude the section we would like to draw your attention to the fact that we can prove some theorems by two methods, direct and indirect.

In the *direct proof* we establish the truth of the proposition being proved by establishing the direct connection between this proposition and those proved before.

In the *indirect proof* we establish the fact that should we raise doubts as to validity of the proposition being proved and regard it as false we would arrive at a contradiction either with the conditions, or with propositions proved before. For this reason the indirect proof is also known as *reductio ad absurdum*.

In the above we made use mainly of the direct proof. Let's cite now some examples of reductio ad absurdum.

As a first example we shall cite the third criterion for the equality of triangles. The approved textbook states that it is inconvenient to prove this criterion by superposing triangles since we know nothing about the equality of the angles. However, using the reductio ad absurdum method it is possible to use superposition to prove this criterion, as well.

Let ABC and $A'B'C'$ be the given triangles (Fig. 22) where $BC = B'C'$, $CA = C'A'$, $AB = A'B'$. To prove the point let's superpose $\triangle A'B'C'$ on $\triangle ABC$ so that the side $A'B'$ coincides with AB.

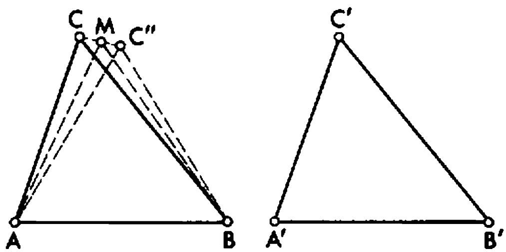

Fig. 22

As we know nothing about the equality of the angles we have no right to assert that the points C and C' will coincide. Let's assume therefore that it will occupy the position C'' Let's connect the points C and C'' The triangle ACC'' is isosceles (to comply to the condition $AC'' = AC$); $\triangle BCC''$ is isosceles, too (from the condition $BC'' = BC$). The altitude AM of the isosceles triangle ACC'' will pass through the point M – the middle of the side CC'' (for in an isosceles triangle the mèdian and the altitude coincide). The altitude BM of the isosceles triangle BCC'' will also pass through the middle M of the side CC'' Thus, we see that two perpendiculars AM and BM have been erected from the point M to the line CC'' These perpendiculars cannot coincide for it would mean that the points A, B and M belong to the same line which is impossible because the points C and C'' (and therefore the entire segment) lie to the same side of the line AB.

Hence, we have arrived at the conclusion that should we assume that the points C and C' will not coincide we shall have to admit that it is possible to erect two different perpendiculars from the same point M to the line CC'' But this is in contradiction with the properties of the perpendicular established before.

Consequently, on superposition of the triangle the point C' must coincide with the point C and we shall find that $\triangle ABC = \triangle A'B'C'$

As a second example we shall take the proof of the assumption expressed before that if two bisectors of a triangle are equal that triangle is an isosceles one.

Suppose we have $\triangle ABC$ and its bisectors AM and BN (Fig. 23).

Let's write down the theorem.

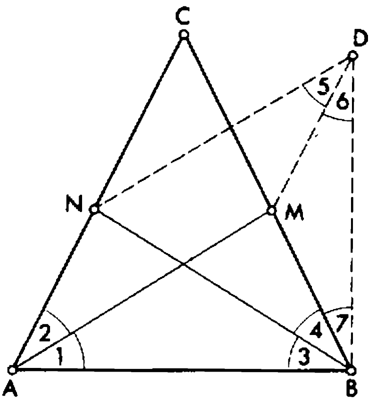

Fig. 23

Given: in $\triangle ABC$ $\angle CAM = \angle BAM$; $\angle CBN = \angle ABN$ and $AM = BN$.

To be proved: $AC = BC$.

We shall use the reductio ad absurdum method of proof. Let's suppose that the triangle is not isosceles and that, to be definite, $AC > BC$. If it is so, then $\angle ABC > \angle CAB$, as well. Ascribing numbers to the angles as shown in the drawing we obtain $\angle 3 > \angle 1$. Next compare $\triangle ABM$ and $\triangle ABN$; AB is common to both, $AM = BN$ from the condition, but the angles between the respective equal sides are not equal. Therefore, the side lying opposite the greater angle will be greater, too, i. e. $AN > BM$. Let's draw the segment ND equal and parallel to AM through the point N. In that case the quadrilateral $AMDN$ will be a parallelogram and therefore $MD = AN$ and $\angle 5 = \angle 2$. Connecting B with D we obtain $\triangle BDN$ which will be an *isosceles* one for $ND = AM = BN$. On the other hand, in $\triangle BDM$ the side $MD = AN$, but $AN > BM$ and therefore $MD > BM$, whence $\angle 7 > \angle 6$. At the same time, $\angle 4 > \angle 5$ because $\angle 5 = \angle 2 = \angle 1$ and $\angle 4 = \angle 3$, but $\angle 3 > \angle 1$. Should we add

up the inequalities $\angle 7 > \angle 6$ and $\angle 4 > \angle 5$ term by term we would obtain: $\angle 4 + \angle 7 > \angle 5 + \angle 6$, i. e. $\angle BDN > \angle DBN$. We have found that the angles at the base of an isosceles triangle BDN are not equal. The resulting contradiction makes us reject the assumption that $AC > BC$. In the same way we could have refuted the assumption that $BC > AC$. Hence, $AC = BC$.

The above examples make the point of the nature of the reductio ad absurdum proof sufficiently clear. We usually resort to this method when in the process of the search for arguments it is established that it is difficult, or sometimes even impossible, to find a direct proof.

It is customary in such cases to turn to a proposition contrary to that which it is required to prove and to try by means of analysis to find such a set of judgements that would lead to a proposition contradicting some previously established proposition. Such were the contradictory propositions we arrived at in the two latter examples: in the first we arrived at the conclusion that two perpendiculars to a single line can be erected from the same point and in the second that the angles at the base of an isosceles triangle are not equal.

§ 4. What Propositions May Be Accepted Without Proof?

1. Let's now answer the last question formulated in the introduction: what propositions in geometry may be accepted without proof?

This question seems a very simple one at first glance. Anyone will say that axioms may be accepted without proof; as such one should take propositions the truth of which has been tested repeatedly in practice and is beyond any doubt. However, when we try and choose such propositions it turns out to be not so simple practically.

At present we know many geometrical propositions that have been tested in practice so often as to exclude any doubt as to their validity. But it by no means follows from here that all those propositions should be accepted as axioms. For instance, we do not doubt that only one straight line can be drawn through two points; that one and only one perpendicular to a line can be drawn through a given point; that the sum of two sides of a triangle is larger than its third side; that two segments each equal to a third are equal between themselves; that the distance between two

parallel lines is everywhere the same, etc. Clearly, the list of such propositions could be lengthened many times over again. Why not accept all these propositions as axioms? This would make the presentation of geometry much more simple, many proofs would have become superfluous, etc.

However, the progress of geometry did not take this course: on the contrary, the geometers strived to cut the number of axioms down to the minimum and to deduce the remaining material of geometry from this small number of fundamental truths.

Why did they take this seemingly more difficult and complicated course of constructing the system of geometrical knowledge?

The desire to build geometry on the basis of a minimum number of axioms is due to a number of causes. Firstly, as the number of axioms decreases the importance of each rises: we should not forget that axioms must contain all the future geometry that is to be deduced from them. Therefore, the less the number of axioms the more general, more profound and more important are the properties of spatial forms that are reflected by each separate axiom.

Another important reason to minimize the number of axioms is that it is much easier to test the truth of a small number of axioms and to see how the conditions set for the totality of axioms (we shall deal with them below) are fulfilled.

2. Thus, we are faced with the problem of choosing the minimum number of the most fundamental and most important propositions of geometry which we shall accept as axioms. What is to guide us in this choice? We should first of all keep in mind that we should not choose axioms considering them one by one, independently of other axioms. We should accept not a single axiom, but an *entire system of axioms*, because it is only such a system that can correctly reflect the real properties and interrelations of the main spatial forms of the material world.

Naturally, only repeatedly tested facts that reflect the most general and fundamental laws of spatial forms may be included in such a system.

Next, in accepting such a system of axioms we must make sure that it will not include contradictory propositions for such propositions cannot all be true at the same time. For example, it is not permissible for the system to include simultaneously the axioms: "Not more than one line parallel to the given straight line can be drawn through the given point" and "No line parallel to the given line can be drawn through the given point"

Not only the axioms themselves should not be contradictory, but among the conclusions from them there should not be two

propositions contradicting each other. This main requirement to be satisfied by the system of axioms is termed the *condition of consistency.*

But at the same time, we should see to it that our system of axioms does not include a proposition provable with the aid of the other axioms. This requirement will be clear if we remember that we strive to minimize our system, i. e. make it contain the least number of propositions which are taken without proof. If a proposition can be proved on the basis of other axioms, then it is a theorem rather than an axiom and should not be included into the system of axioms. The requirement that the axiom should not be a consequence of the other axioms is termed the *condition of independence.*

However, in striving to minimize our system of axioms we should not go too far and exclude such propositions from it that we shall unavoidably have to use as a basis in the presentation of geometry.

This is the third condition that a system of axioms should fulfil – *the condition of completeness of the system.* To be exact, this condition may be formulated as follows: if a system is *incomplete* it is always possible to add a new proposition to it (clearly a proposition containing the same basic concepts as other axioms) that would be independent of the other theorems and would not contradict them. In case of a *complete* system every new proposition containing the same concepts as in axioms will either be a conclusion drawn from these axioms or will contradict them.

3. To gain a clearer understanding of the conditions of completeness, independence and consistency for a system of axioms one could turn to a simple example which, although not an exact reflection of geometrical relations, presents a fair analogy with them.

Let's consider a system of equations in three unknowns. We shall regard each of the three unknowns as some "concept" requiring definition and each equation as a sort of an "axiom" with the aid of which the relationship between the "concepts" is established.

So let's suppose we have a system

$$2x - y - 2z = 3$$
$$x + y + 4z = 6$$

Can the unknowns x, y and z be determined from this system? No, they cannot, since the equations are fewer than the unknowns. The system does not comply with the *condition of completeness.*

Next we shall try to modify the system by supplementing it with yet another equation:

$$\begin{gathered} 2x - y - 2z = 3 \\ x + y + 4z = 6 \\ 3x + 3y + 12z = 18 \end{gathered}$$

After careful consideration of the modified system we conclude that the addition of the new equation did not change the situation because the third equation is simply a consequence of the second and does not introduce any new relations. The system violates the *condition of independence.*

Let's change the third equation and consider this system:

$$\begin{gathered} 2x - y - 2z = 3 \\ x + y + 4z = 6 \\ 3x + 3y + 12z = 15 \end{gathered}$$

Again it may easily be seen that this system, too, cannot be used to determine the unknowns.

Indeed, dividing both parts of the last equation by 3 we obtain the equation

$$x + y + 4z = 5$$

On the other hand, the second equation yields:

$$x + y + 4z = 6$$

Which of the two equations should be believed? Clearly, we have here an *inconsistent* system which is also useless for the determination of the unknowns.

Should we at last consider the system

$$2x - y - 2z = 3$$

$$x + y + 4z = 6$$

$$2x + y + 5z = 8$$

we would easily conclude that it has a unique solution ($x = 5$, $y = 13$, $z = -3$), that it is consistent, independent and complete. Should the fourth equation be added to the system it would turn out to be either a consequence of the given three, or would contradict them.

4. We see from here that the choice of axioms to serve as a basis of geometry is far from arbitrary, but subject to very serious

requirements. The work of setting up the necessary system of axioms was started as far back as the end of the past century and although the scientists have done a great deal in this direction it still cannot be regarded as completed. This is because in the course of the revision of the available system of axioms scientists discover from time to time the presence of superfluous, i. e. "dependent" axioms, that are the consequence of more simple and more general axioms, and therefore complicated propositions containing numerous conditions are replaced by axioms with fewer conditions, etc. This research is of great interest to science for it aims at establishing what are the most general, profound and important properties of spatial forms that determine the entire context of geometry.

In order to throw some light on the system of axioms of modern geometry let us turn first to the presentation of geometry at school and see what axioms it is built on, what axioms it lacks. We shall limit ourselves to the axioms of plane geometry.

The presentation of geometry at school starts with the explanation of the basic concepts of geometry: the body, the surface, the line, the point. Next the straight line is selected from lines of all sorts and the plane from the surfaces. The first axioms of the school course establish the relations between the point, the straight line and the plane. These are *axioms of connection* – the first group in the complete system of axioms of geometry.

The axioms of this group establish how the main geometric elements are connected with each other: how many points determine a straight line, or a plane, what are the conditions for the straight line to belong to a plane, etc.

Out of the axioms of connection only two are mentioned in the school course:

(1) *One and only one straight line passes through any two points.*

(2) *If two points of a straight line lie in a plane the entire line lies in the plane.*

At the same time we, consciously, or instinctively, make constant use of other axioms of connection, as well, from whose number the following are also necessary to substantiate plane geometry:

(3) *On every straight line there are at least two points.* The requirement of this axiom is, as we see, quite limited. However, in future we shall be able to prove with the aid of axioms of order the existence of an infinite number of points on a straight line.

(4) *There exist at least three points in the plane not lying on one*

straight line. This axiom, too, contains the minimum requirement on the basis of which the existence of an unlimited number of points in a plane can be proved in future.

5. We shall now turn to the second group of axioms that is totally absent from the school course, although one has to use them time and again. The axioms of the second group are termed *axioms of order*. These axioms describe the laws that govern the mutual position of the points on a straight line and that of the points and straight lines in a plane. We often use these axioms

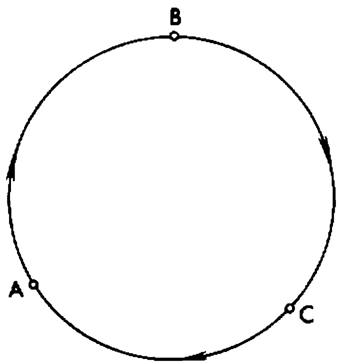

Fig. 24

although in an implicit form. If, for instance, we have to prolong a segment we do it knowing that it is always possible to prolong a segment both ways.

If we connect two points lying to two different sides of a straight line we are confident that the segment thus obtained will intersect the line. We relied on this fact, for instance, while proving the theorem that triangles with two equal sides and an equal angle lying opposite one of the sides are equal (see Fig. 12). One more example: we are confident that the bisector of an interior angle of a triangle will not fail to intersect the opposite side.

Undoubtedly, all those are quite obvious facts, but still they speak of the existence of some basic properties of the geometrical figures that we constantly use and that for this reason should find their place among axioms.

The axioms that determine the position of the points on a straight line are bound up with basic concepts such as "to precede" and "to follow" and are formulated as follows:

(1) Any one of the two points lying on one straight line may

·e defined as the preceding one, in that case the second be the following.

(2) *If A, B, C are points on the same straight line and if A precedes B, B precedes C, then A precedes C.*

Already these two axioms define rather clearly the peculiarities of the straight line which are not characteristic of all lines. Let's take, for example, a circle (Fig. 24) and moving clockwise along it mark in turn the points A, B, C; we shall then see that on the circle point A precedes point B, point B precedes point C and point C again precedes point A. When the position of the points A, B and C on the straight line is as has been stated above, we say that B lies *between* A and C (Fig. 25).

Fig. 25

(3) *Between any two points of a straight line there is alway. another point of the same line.*

Applying this axiom in turn to two points on the straight line (they exist by force of the second axiom of connection), next to each of the intervals thus obtained, etc., we find that between any two points of a straight line there is an infinite number of points of the same line.

The part of a straight line on which lie two of its points and all the points between them is termed a *segment.*

(4) *Every point of a straight line has both a preceding and a following point.*

As a consequence of this axiom a segment of a straight line can be prolonged both ways. Another consequence is that there is no point on a straight line that would precede, or follow, all the other points, e. that a *straight line has no ends.*

The part of a straight line to which belong the given point and all its preceding, or following points, is termed a *ray* or *half-line.*

The mutual position of points and straight lines in a plane is determined by the following axiom termed the "Pasch's axiom" after the German mathematician who first formulated it.

(5) *If there exist three points not lying on one line, then the line lying in the same plane and not passing through these points and intersecting a segment joining these points intersects one and only one other segment* (Fig. 26).

This axiom is used to prove the theorem on a straight line

dividing a plane into two half-planes. Let's cite the proof of this theorem as an example of rigorous proof that relies only on axioms and on theorems proved before. We shall formulate the theorem as follows.

A straight line lying in a plane divides all the points of the plane not lying on the line into two classes such that points of one class determine a segment not intersecting the line and the points of distinct classes determine a segment intersecting the line.

In the course of the proof we shall make use of some special designations which ought to be remembered. $\subset$ is the sign of *belonging*:

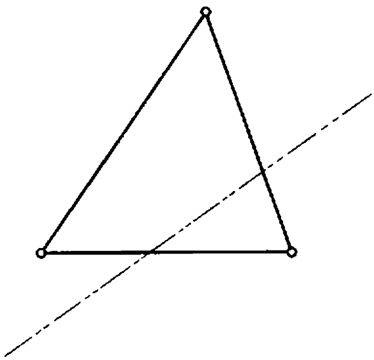

Fig. 26

$A \subset a$ – “the point A belongs to the straight line a” $\times$ is the *intersection* sign; $AB \times a$ – “the section AB intersects the straight line a” A bar placed above some relation means its *negation*: $\overline{A \subset a}$ – “the point A does not belong to the straight line a”

is the symbol of drawing a *conclusion* – “therefore” Having adopted this notation we now turn to the proof of the theorem. First of all note that *if the three points lie on the same straight line, for them, too, holds a proposition similar to Pasch's-axiom: a line intersecting one of the three segments determined by these three points intersects one and only one other segment.* This proposition may easily be proved on the basis of axioms dealing with the position of points on a straight line.

Indeed, if the points A, B and C lie on a single straight line and point B lies between the points A and C all the points of the segments AB and BC belong to the segment AC and every point of the segment AC (excluding B) belongs either to AB, or to BC.

Therefore, a line intersecting AB, or BC, will perforce intersect AC, and a line intersecting AC will intersect either AB, or BC.

Suppose now we have the straight line l lying in a plane. We must prove the following:

(1) It is possible with the aid of this line l to divide the points of the plane that do not lie on this line into classes.

(2) There can be two and only two classes.

(3) The classes have the properties that are stated in the theorem.

To establish this fact let's take the point A not lying on the line l (Fig. 27) and adopt the following conditions:

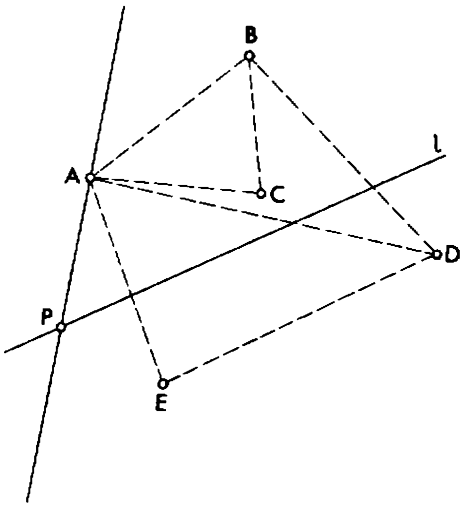

Fig. 27

(a) point A belongs to the *first* class (let's denote it K_1);

(b) a point not lying on l belongs to the *first* class if it, together with the point A, determines a segment that does not intersect l;

(c) a point not lying on l belongs to the *second* class (let's denote it K_2) if it, together with the point A, determines a segment that intersects l.

It may easily be seen that there are points of both classes. Let's take a point P on the line l and draw a straight line PA. The ray with the vertex P containing the point A contains only points of the first class since the point of intersection P lies outside the segments determined by the point A and the other points of the ray. The opposite ray with the same vertex contains only

points of the second class since the point of intersection P lies inside all the segments determined by the point A and the points of this ray. Connecting A with any point of the line l we shall obtain an infinite number of straight lines containing the points of the first and the second class.

There can be *only two* classes since we can state *only two* propositions in respect of any segment that connects A with a point not lying on l: either the segment intersects l, or it does not – there can be no third possibility.

Lastly, let's show that the classes K_1 and K_2 satisfy the conditions of the theorem. Let's consider the following cases.

(1) Both points belong to the first class: $B \subset K_1$, $C \subset K_1$. Since $B \subset K_1$ then $\overline{AB \times l}$; since $C \subset K_1$ then $\overline{AC \times l}$. by force of Pasch's axiom $\overline{BC \times l}$.

(2) Both points belong to the second class: $D \subset K_2$ and $E \subset K_2$. Since $D \subset K_2$ then $AD \times l$; since $E \subset K_2$ then $AE \times l$. by force of Pasch's axiom $\overline{DE \times l}$.

(3) The points belong to different classes: $B \subset K_1$; $D \subset K_2$. Since $B \subset K_1$ then $\overline{AB \times l}$; since $D \subset K_2$ then $AD \times l$. by force of Pasch's axiom $BD \times l$.

The theorem has been proved.

The part of a plane to which belong all the points of one class is termed a *half-plane*.

Note that the theorem can be proved without the use of a drawing. The drawing only helps to follow the course of the reasoning and to memorize the obtained relationships. This, by the way, is true of any sufficiently rigorous proof.

6. The following, third, group of axioms deals with the concept of *equality*. In the school geometry course the equality of figures in a plane is established by superposition of one figure on another.

The approved geometry textbook treats this problem as follows: "Geometrical figures may be moved about in space without changing either their shape or size. Two geometrical figures are termed equal if by moving one of them in space it can be superposed on the other so that both figures will coincide in all their parts."

At first glance this definition of equality seems to be quite comprehensible, but if one considers it carefully a certain logical circle may easily be found in it. Indeed, to establish the equality of figures we have to superpose them, and to do this we must *move* one figure in space, the figure remaining *unchanged* in the process. But what does "remaining unchanged" mean? It means that the figure all the time remains equal to its original image. Thus,

it comes about that we define the concept of equality by moving an "unchanged figure" and at the same time we define the concept "unchanged figure" by means of "equality."

Therefore there seems to be much more sense in establishing the equality of figures by means of a group of axioms dealing with the equality of segments, angles and triangles.

The axioms that establish the properties of the equality of segments are the following:

(1) *One and only one segment equal to the given segment can be marked off on a given line in a given direction from a given point.*

(2) *Every segment is equal to itself. If the first segment is equal to the second, the second is equal to the first. Two segments each equal to a third are equal to each other*

(3) *If A, B and C lie on the same straight line and A', B' and C', too, lie on the same straight line and if $AB = A'B'$, $BC = B'C'$ then $AC = A'C'$, too.*

In other words, should equal segments be added to equal ones the sums would be equal, as well.

There are quite similar axioms for the angles.

(4) *One and only one angle equal to the given angle can be built at the given ray in the given half-plane.*

(5) *Every angle is equal to itself. If the first angle is equal to the second, the second is equal to the first. If two angles are each equal to a third they are equal to each other.*

(6) *If a, b and c are rays with a common vertex, a', b' and c' are other rays with a common vertex and if $\angle ab = \angle a'b'$, $\angle bc = \angle b'c'$, then $\angle ac = \angle a'c'$, too.*

In other words, should equal angles be added to the equal ones the sums would be equal, too.

Lastly, one more axiom of the third group is introduced to substantiate the equality of triangles.

(7) *If two sides and the angle between them of one triangle are equal to the respective sides and the angle between them of the second, the other respective angles of these triangles are equal, too.* If, for example, we have $\triangle ABC$ and $\triangle A'B'C'$ with $AB = A'B'$, $AC = A'C'$ and $\angle A = \angle A'$, then $\angle B = \angle B'$ and $\angle C = \angle C'$, as well.

These seven axioms are used first to prove the main criteria of the equality of triangles, to be followed by all the theorems dealing with the equality of figures that are based on those criteria. Now no more need arises for the method of superposition, it becomes superfluous.

Let's see, for instance, how the first criterion of the equality of triangles is proved.

Suppose $\triangle ABC$ and $\triangle A'B'C'$ are given (Fig. 28) with $AB = A'B'$, $AC = A'C'$ and $\angle A = \angle A'$. It is required to prove the equality of all the other elements of the triangles. From axiom (7) we obtain immediately that $\angle B = \angle B'$ and $\angle C = \angle C'$ It remains for us to show that $BC = B'C'$

Suppose that $BC \neq B'C'$ Then we mark off $B'C'' = BC$ on the side $B'C'$ from the point B' Consider $\triangle ABC$ and $\triangle A'B'C''$

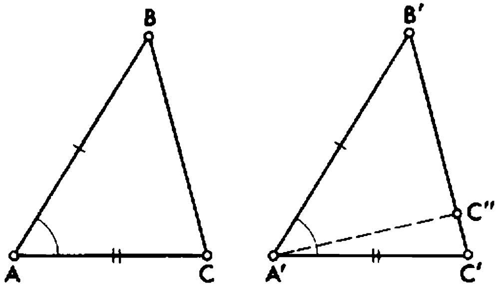

Fig. 28

They have $AB = A'B'$, $BC = BC''$ and $\angle B = \angle B'$ Then in compliance with axiom (7) $\angle B'A'C'' = \angle A$, too. But two angles equal to a third are themselves equal, therefore $\angle B'A'C'' = \angle B'A'C'$ We see that two different angles each equal to the same angle A have been built at the ray $A'B'$ in the same half-plane and this contradicts axiom (4). Hence, if we reject the supposition $BC \neq B'C'$ we shall obtain $BC = B'C'$

The proof of other theorems dealing with the equality of figures is similar.

7. As the presentation of the elementary geometry proceeds, the need arises for yet another group of axioms, i. e. the *axioms of continuity*, to be introduced. The problems of intersection of a line and a circle and of intersection of circles are closely related to the axioms of this group. These are the problems upon which all the geometrical constructions made with the aid of compasses and a ruler are based. This fact speaks for the enormous importance of the axioms of continuity. Moreover, the entire theory of measurement of geometrical quantities is built around the axioms of continuity.

The group of axioms of continuity includes the following two axioms:

(1) **The Archimedean axiom.** *If two segments are given, one of them greater than the other, then by repeating the smaller segment a sufficiently large number of times we can always obtain a sum that exceeds the larger segment.* In short, if $\bar{a}$ and $\bar{b}$ are two segments and $\bar{a} > \bar{b}$ there exists an integer n such that $n\bar{b} > \bar{a}$.

The Archimedean axiom found its place in the approved textbook, as well, namely, in the chapter on the measurement of segments. The above-mentioned method of finding a common unit of length of two segments by means of successive laying off is based on the Archimedean axiom. Indeed, this method entails the laying off of the small segment on the larger one, and the

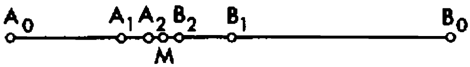

Fig. 29

Archimedean axiom assures us that with such a procedure the sum of small segments will ultimately cover the large segment.

We conclude directly from the Archimedean axiom that if the interval $\bar{a}$ is greater than the interval $\bar{b}$ there always exists an integer n such that $\frac{\bar{a}}{n} < \bar{b}$.

The second of the continuity axioms bears the name of *Cantor*, or the *nested intervals axiom*. Here is how it reads:

(2) *If there is a system of intervals wherein each succeeding interval is inside the preceding one and if within this system there can always be found an interval that is smaller than any given one then there is a single point lying inside all these intervals.*

In order to illustrate the way in which the Cantor axiom is used let's consider the following example. Let's take the interval A_0B_0 (Fig. 29), denote its mid-point by B_1 and find the middle of the interval A_0B_1 which we shall denote by A_1. Next we take the middle of A_1B_1, denote it by B_2 and find the middle of the interval A_1B_2 which we denote by A_2. Next we take the middle of A_2B_2 which we denote by B_3, find the middle of A_2B_3 and denote it by A_3. Next we take the middle of A_3B_3, etc.* The intervals A_0B_0, A_1B_1, A_2B_2, A_3B_3, constitute a system of nested intervals. Indeed, each succeeding interval is inside the preceding one and is equal to $^1/_4$ of it. Thus, the length of the interval A_1B_1 is equal

* There is no room in the drawing for this interval A_3B_3, so it has to be imagined.

to $^1/_4\ A_0B_0$, the length of $A_2B_2 = {}^1/_{16}\ A_0B_0$, $A_3B_3 = {}^1/_{64}\ A_0B_0$, and in general $A_nB_n = \dfrac{A_0B_0}{4^n}$

It follows from the Archimedean axiom that the length $\dfrac{A_0B_0}{4^n}$ obtained in this manner can be smaller than any given segment for a sufficiently large n. Hence, all the conditions of the axiom are satisfied and there is a single point lying inside the entire system of segments. This point may easily be shown. Indeed, if we take point M on $^1/_3$ of the segment A_0B_0, i. e. so that $A_0M = {}^1/_3\ A_0B_0$ this will be the point we are looking for. Indeed, should we take point A_0 as the origin of the number axis and assume the segment A_0B_0 to be unity the following numerical values would then correspond to the points A_1, A_2, A_3, A_n:

$$\frac{1}{4};\ \frac{1}{4}+\frac{1}{4^2}=\frac{5}{16};\ \frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}=\frac{21}{64};\qquad \frac{1+4+4^2+\ ..+4^{n-1}}{4^n}$$

Each of these fractions is less than $/_3$.

Indeed, if we subtract unity from the denominator of each of the fractions, the fraction will become larger and exactly equal to $^1/_3$:

$$\frac{1+4+4^2+\ \dots\ +4^{n-}}{4^{n-1}} =$$

$$= \frac{1+4+4^2+\ \dots\ +4^{n-}}{(4-1)(1+4+4^2+\qquad +4^{n-1})} = \frac{1}{3}*$$

On the other hand, the corresponding numerical values for the points B_1, B_2, B_3, B_n are

$$\frac{1}{2};\ \frac{1}{2}-\frac{1}{8}=\frac{3}{8};\ \frac{1}{2}-\frac{1}{8}-\frac{1}{32}=\frac{11}{32};\qquad \frac{1}{2}-\frac{1}{8}-\frac{1}{32}-\qquad \frac{1}{2^{2n+1}}$$

The numerical value corresponding to the point B_1 may also be written in the following form:

$$\frac{1}{2}-\left(\frac{1}{4}-\frac{1}{8}\right)-\left(\frac{1}{16}-\frac{1}{32}\right)-\qquad -\left(\frac{1}{2^{2n}}-\frac{1}{2^{2n+1}}\right)=$$

$$=\frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{1}{16}+\frac{1}{32}-\qquad -\frac{1}{2^{2n}}+\frac{1}{2^{2n+1}}$$

* Here we make use of the formula
$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \dots + ab^{n-2} + b^{n-1})$

Should we add these numbers up we would obtain

$$\frac{2^{2n} - 2^{2n-1} + 2^{2n-2} \quad \ldots - 2^3 + 2^2 - 2 + 1}{2^{2n+1}}$$

It may easily be obtained from here that each numerical value that corresponds to the points B_1, B_2, $\quad B_n$ is *greater* than $^1/_3$. Adding unity to the denominator we thereby make the fraction smaller obtaining

$$\frac{2^{2n} - 2^{2n-1} + 2^{2n-2} - \ldots - 2^3 + 2^2 - 2 + 1}{2^{2n+1} + 1} =$$

$$= \frac{2^{2n} - 2^{2n-1} + 2^{2n-2} - \ldots - 2^3 + 2^2 - 2 + 1}{(2 + 1)(2^{2n} - 2^{2n-1} + 2^{2n-2} \quad - 2^3 + 2^2 - 2 + 1)} = \frac{1}{3}*$$

Hence all the numerical values that correspond to the points B_1, B_2, B_3, .., B_n, are greater than $^1/_3$. It follows from here that the point M with the corresponding numerical value equal to $^1/_3$ is inside each of the segments A_1B_1, A_2B_2, A_3B_3, ., A_nB_n, Therefore, this is the unique point determined by the sequence of these segments.

Let's now turn to the proof of the basic theorem on the intersection of a straight line and a circle.

Let's recall that a circle is determined by its centre and its radius. The points of the plane the distance from which to the centre is less than the radius are termed *interior* points in respect to the circle; the points the distance from which to the centre is greater than the radius are termed *exterior* in respect to the circle.

The basic theorem is formulated as follows:

A segment that connects an interior point in respect to the circle with an exterior one has one and only one common point with the circle.

Suppose we have a circle with the centre at point O and the radius r; A is an interior point ($OA < r$), B is an exterior point ($OB > r$) (Fig. 30). Let's prove, to begin with, that if there is a point M on AB the distance from which to point O is equal to the radius, that point will be unique. Indeed, if such a point M

* Here we make use of the formula

$$a^{2n+1} + b^{2n+1} = (a + b)(a^{2n} - a^{2n-1}b + a^{2n-2}b^2 - \ldots - ab^{2n-1} + b^{2n})$$

exists, there should also exist a point M' symmetrical with M in respect to the perpendicular dropped from O to AB with $M'O = MO = r$. By force of the properties of inclined lines drawn from a point to the straight line AB all the interior points of the segment $M'M$ will be interior points of the circle, as well, and all the exterior points of the segment $M'M$ will be exterior points of the circle. Therefore, the point A must always lie between the points M' and M and only one point M may lie on the segment AB.

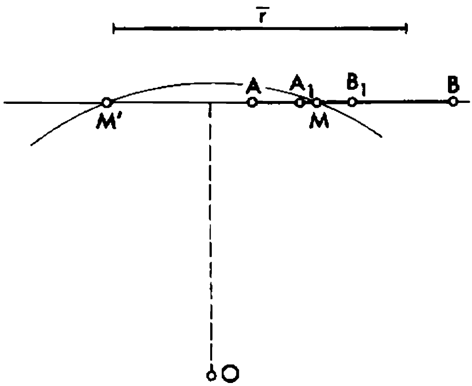

Fig. 30

Having established this fact let's divide the segment AB in two and compare the distance of the point thus obtained from the centre with the radius. If this distance turns out to be equal to the radius, the theorem will have been proven. If this distance turns out to be less than the radius, the point will be an interior one and we shall denote it A_1. If this distance turns out to be greater than the radius, the point will be an exterior one and we shall denote it B_1.

Next we take the middle of the segment A_1B (or AB_1) in respect to which there are again three possible cases: either the distance from it to the centre is equal to the radius and in that case the theorem will have been proven, or it is less than the radius – in that case we denote this point A with a corresponding numerical index, or it is greater than the radius – in that case we denote this point B with a corresponding numerical index. Continuing this process indefinitely we find that either the distance of one of such points from the centre is equal to the radius and this proves the theorem, or all the points denoted by the

letters A_1, A_2, , A_n, will be interior and all the points denoted by the letters B_1, B_2, B_n, will be exterior. But in this latter case we obtain a system of segments satisfying the conditions of the Cantor axiom for each of the succeeding segments lies inside the preceding one and the length of each succeeding segment is half that of the preceding segment. Therefore, there exists a unique point that lies inside all those segments. Since it lies between all the interior and all the exterior points of the segment it can be neither an interior, nor an exterior one; therefore, it is a point of the circle.

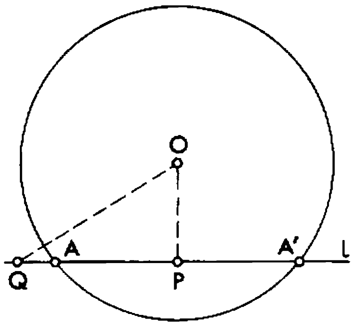

Fig. 31

It follows from this theorem, in particular, that if the distance from the straight line to the centre of a circle is less than the radius, this line will have two and only two points in common with the circle. Indeed, let O be the centre and r the radius of the circle (Fig. 31). The distance OP from the centre to the line l is less than the radius; therefore, P is an interior point. Let's mark off the segment $PQ = r$ on the line l from the point P

Since the hypotenuse OQ of the right triangle OPQ is greater than the leg $PQ = r$, it follows that $OQ > r$ and, therefore, Q is an exterior point. According to the theorem just proved the segment PQ has one point A in common with the circle. The second common point A' is symmetrical with A in respect of the perpendicular OP. Since all the interior points of the segment AA' are also the interior points of the circle and all the exterior points are exterior in respect of the same circle, the line l has no other common points with the circle.

The propositions similar to the Archimedean and the Cantor axioms may be proved for the arcs of a circle as well, e. it is possible to prove that:

(1) *By repeating a given arc a sufficiently large number of times we can obtain an arc greater than any predetermined arc.*

(2) *If we have a system of arcs in which each succeeding arc lies inside the preceding one and if it is always possible to find an arc in the system that is smaller than any given arc, there is a point that lies inside all these arcs.*

On the basis of these propositions one may easily prove the basic theorem on the intersection of circles:

If A is the interior and B the exterior point in respect of the given circle, then the arc of any other circle that connects A and B has one and only one common point with the given circle.

The proof of this theorem is quite similar to that of the theorem on the intersection of a straight line and a circle.

8. The last, fifth, group of geometrical axioms deals with the concept of *parallelism* and consists of only one axiom:

Only one line can be drawn parallel to a given line through a given point not on this line.

The propositions based on this axiom are widely known and we shall not stop to consider them.

The system of axioms discussed above gives an idea of the totality of propositions taken without proof that can make the basis of geometry. But it should be noted that aiming at the simplification of presentation we made no attempts to minimize the system. The number of those axioms could be brought down still further. For instance, two axioms – those of Archimedes and Cantor – could be replaced by one, the so-called Dedekind axiom. The conditions of the axioms could be made less strict. For example, it would be possible to refute the requirement that the straight line in Pasch's axiom which intersects one of the sides of a triangle should intersect one and *only* one other side. Actually, it is possible to retain the only requirement that a straight line which intersects one of the sides of a triangle should intersect another side and to prove that there will be *only* one such side. In the same way in the formulation of Cantor's axiom the requirement that the point determined by the system of nested intervals be unique may be refuted. The uniqueness of this point, too, can be proved. All this would, however, make the presentation more elaborate and complicated.

Let's summarize in conclusion the themes we have discussed in this booklet.

(1) We defined geometry as a science dealing with the spatial forms of the material world.

(2) We obtained initial knowledge of the spatial forms by way of induction, i. e. through repeated observations and experiments.

(3) We formulated the most profound and most general spatial properties of things in the form of a system of fundamental propositions – axioms.

(4) A system of axioms will correctly reflect the real spatial properties only if it satisfies the conditions of completeness, independence and consistency.

(5) With the exception of axioms all the other propositions of geometry – theorems – are obtained by way of deduction from the axioms and from the theorems proved before. This system of deduction is called proof.

(6) For the proof to be correct, i. e. the validity of the theorem being proved to be beyond doubt, it must be built on correct judgements and must be free from errors. The correctness of a proof depends on: (1) an accurate correct formulation of the proposition being proved, (2) the choice of the necessary and true arguments and (3) rigorous adherence to the rules of logic in the course of the proof.

TO THE READER

Mir Publishers welcome your comments on the content, translation and design of this book.

We would also be pleased to receive any proposals you care to make about our future publications.

Our address is:

USSR, 129820, Moscow I-110, GSP
Pervy Rizhsky Pereulok, 2
Mir Publishers

Printed in the Union of Soviet Socialist Republics